

EXHIBIT 76

Confidential.

**UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION**

-----	X	
	:	
Richard Kadrey, Sarah Silverman, Christopher	:	
Golden, Ta-Nehisi Coates, Junot Díaz, Andrew	:	
Sean Greer, David Henry Hwang, Matthew Klam,	:	
Laura Lippman, Rachel Louise Snyder,	:	
Jacqueline Woodson, Lysa TerKeurst,	:	
and Christopher Farnsworth,	:	
	:	Case No. 3:23-cv-03417-VC
Plaintiffs,	:	
	:	
v.	:	
	:	
Meta Platforms, Inc., a Delaware corporation;	:	
	:	
Defendant.	:	
	:	
-----	X	

EXPERT REPORT OF DANIEL F. SPULBER

JANUARY 10, 2025

190. Regarding torrenting of copyrighted works, Marybeth Peters, the Register of Copyrights at the U.S. Copyright Office as of 2003, has stated that “the law is unambiguous. Using peer-to-peer networks to copy or distribute copyrighted works without permission is infringement, and copyright owners have every right to invoke the power of the courts to combat such activity. Every court that has addressed the issue agrees.”⁴⁶³
191. From an economic perspective, downloading illegal, pirated copies of copyrighted works and then redistributing those same pirated copies en masse (through torrenting and seeding) to other users risks reducing demand for those works in the legal market. Such conduct causes concrete harm to authors by diminishing demand for their copyrighted works.

C. Meta Risks Depriving Plaintiffs of Revenues by Using Copyrighted Books as Inputs to Allow the Creation of Works that Compete with Plaintiffs’ Works.

192. Meta risks depriving Plaintiffs of revenues by using copyrighted books as inputs to allow the creation of works that could compete with Plaintiffs’ works. Meta’s Llama models, trained on Plaintiffs’ works, are capable of outputting or substantially assisting in outputting fiction and nonfiction works that will compete with and decrease demand for Plaintiffs’ works. These works risk depriving Plaintiffs of revenues that would have been obtained in the market for books and other subsidiary rights. These works would be economic substitutes for Plaintiffs’ works that deprive Plaintiffs of revenues, but use the Plaintiffs’ works as inputs.
193. AI LLMs such as Meta’s Llama can generate outputs that are substitute products for authors’ works, thereby decreasing demand for authors’ works. These substitute products generated by AI compete with authors’ works. They may be sold in the market for books and compete directly with authors’ books.
194. In economics, substitute goods are those for which an increase in the price of one good leads to an increase in the demand for the other, and “are often pairs of goods that are used in place of each other[.]”⁴⁶⁴ Substitute goods refer to products that provide consumers with similar but not necessarily identical consumption benefits. Substitute products serve similar purposes for consumers but need

⁴⁶³ *Sen. Hearing 108-920, Before Comm. on the Judiciary, U.S. Senate* (2003) (Statement of Marybeth Peters, Register of Copyrights, U.S. Copyright Office).

⁴⁶⁴ Mankiw, N. Gregory. *Principles of Microeconomics, Fifth Edition*. Mason, Ohio: South-Western Cengage Learning (2009) at 70.

not have identical features.⁴⁶⁵ Because they are substitute products, a decrease in the price of books generated by AI will decrease demand for authors' works.

195. The introduction of AI generated works into the marketplace generates competition that will reduce demand for authors' works. The introduction of AI generated works in the book marketplace works like a reduction in the price of a substitute good (the AI generated works), thus decreasing the demand for authors' works.⁴⁶⁶ The problem is not competition itself, but competition from AI generated works that infringe on authors' copyrighted works.

196. The usage of copyrighted works to train AI LLMs such as Llama would have an adverse effect on creative output by diminishing or eliminating economic incentives to write and by destroying the market for books altogether. Companies that use copyrighted works without compensation to train AI LLMs harm authors of those works in several important ways.

197. First, AI companies such as Meta harm authors by using those copyrighted works to create substitute works that compete with authors' books and diminish authors' income.⁴⁶⁷ There are many examples indicating the prevalence of AI-generated substitute works. WIRED reports that books generated by AI are "[f]looding Amazon"—including, specifically, knockoff books that directly imitate authors' published works by copying their titles and scope of content.⁴⁶⁸ A National Public Radio ("NPR") report finds that there is a growing number of books on Amazon generated by AI, notably "biographies, summaries and even copycat books[.]"⁴⁶⁹ The NPR report quotes Mary Rasenberger, CEO of the Authors Guild, who observes that the problem of scam books has increased: "Every new book seems to have some kind of companion book, some book that's trying to steal sales."⁴⁷⁰ According to the NPR report, "Rasenberger says the publishers posting these books benefit from increasingly more sophisticated AI tools that can generate low-quality 'scam' books quickly."⁴⁷¹ The NPR report further quotes Rasenberger as stating that the problem is outrunning any screening by

⁴⁶⁵ *Id.*

⁴⁶⁶ Baldwin, R. E. and T. Murray. "MFN Tariff Reductions and Developing Country Trade Benefits Under the GSP." *The Economic Journal* 87.345 (1977): 30-46 at 31-33 (discussing the demand effects of an effective price change in the market for imperfect substitute goods).

⁴⁶⁷ Meta researcher Melanie Kambadur confirmed that the downstream applications of LLMs like Llama include as an aid for writing fiction or to output whole poems. Kambadur, Melanie. Deposition (Sept. 17, 2024) ("Kambadur Dep.") 76:7-77:1.

⁴⁶⁸ Knibbs, Kate. "Scammy AI-Generated Book Rewrites Are Flooding Amazon." *WIRED* (Jan. 10, 2024). <<https://www.wired.com/story/scammy-ai-generated-books-flooding-amazon>> (accessed Nov. 4, 2024).

⁴⁶⁹ Limbong, Andrew. "Authors Push Back on the Growing Number of AI 'Scam' Books on Amazon." *NPR* (Mar. 13, 2024). <<https://www.npr.org/2024/03/13/1237888126/growing-number-ai-scam-books-amazon>> (accessed Nov. 4, 2024).

⁴⁷⁰ *Id.*

⁴⁷¹ *Id.*

Amazon: “By the time Amazon finds out about them, they’ve already made some money and they move on to something else[.]”⁴⁷²

198. In one example, an author who had just spent a year researching and writing a nonfiction book about a rare topic discovered another book on the same topic, with the same title, on Amazon three weeks before his own book’s release date.⁴⁷³ Looking further into the situation, the first author, Chris Cowell, discovered no trace of the other purported author online, suspicious five-star Amazon reviews, code snippets that looked like ChatGPT screenshots (the book being a technical software development book), and other indications that the book was written by AI.⁴⁷⁴ The copycat book’s online publisher “listed dozens of books on Amazon on similarly technical topics, each with a different author, an unusual set of disclaimers and matching five-star Amazon reviews from the same handful of India-based reviewers.”⁴⁷⁵
199. There are other examples of AI-generated copycat books, often of low quality, flooding Amazon and affecting demand for authors’ books.⁴⁷⁶ TODAY reports that “[a]rtificial intelligence is making it easier than ever for scammers to impersonate authors and sell fake books online. TODAY’s Savannah Guthrie has firsthand experience with the nefarious trend, as do a number of other authors. Unfortunately, many readers have also unknowingly purchased copycat books online.”⁴⁷⁷
200. The Federal Trade Commission has acknowledged that “examples of fake new songs supposedly from recording artists, as well as new books sold as if authored by humans but in fact reflecting the output of large language models,” have been appearing, conduct that “obviously injures artists and writers[.]”⁴⁷⁸
201. There are more examples of AI-generated books that attempt to substitute for authors’ works. Author Jane Friedman discovered multiple AI-generated books on the publishing industry—her topic of

⁴⁷² *Id.*

⁴⁷³ Oremus, Will. “He Wrote a Book on a Rare Subject. Then a ChatGPT Replica Appeared on Amazon.” *The Washington Post* (May 5, 2023). <<https://www.washingtonpost.com/technology/2023/05/05/ai-spam-websites-books-chatgpt/>> (accessed Oct. 30, 2024).

⁴⁷⁴ *Id.*

⁴⁷⁵ *Id.*

⁴⁷⁶ See, e.g., Callahan, Chrissy and Vicky Nguyen. “Savannah’s Book Became Part of a Copycat Scam. How to Spot AI-Generated, or Fake, Books Online.” *TODAY* (Feb. 27, 2024) <<https://www.today.com/news/ai-generated-books-rcna140666>> (accessed Jan. 9, 2025).

⁴⁷⁷ See, e.g., *id.*

⁴⁷⁸ Atleson, Michael. “Can’t Lose What You Never Had: Claims About Digital Ownership and Creation in the Age of Generative AI.” *Federal Trade Commission* (Aug. 16, 2023). <<https://www.ftc.gov/business-guidance/blog/2023/08/cant-lose-what-you-never-had-claims-about-digital-ownership-creation-age-generative-ai>> (accessed Oct. 30, 2024).

authorship and expertise—that falsely claimed to be authored by her were circulating on Amazon and had been added to her official Goodreads profile.⁴⁷⁹ Amazon initially refused to remove the books, but Friedman took to Twitter and eventually got the platform to reverse that decision, which she credits largely to her visibility and reputation as an already respected author.⁴⁸⁰ In a subsequent tweet, she wondered about the recourse available to authors without the same reputation to trade on.⁴⁸¹ Left without adequate protections, members of authors’ groups have taken to warning each other about the threat of “scammers uploading AI-generated garbage with legitimate authors’ names attached to it.”⁴⁸² One Facebook post in an authors’ group advised:

Because AI allows them to churn out dozens of these impersonations, the volume of those scams has increased dramatically, and the frequency of incorrect author linking has increased with it. Amazon is reportedly working on a fix to prevent these from being linked to an author without authorization, but until that’s ready, it’s a good idea to check your author page periodically to ensure nothing was added to it without your approval. Regardless of Amazon’s fix, it’s a good idea to search for your pen name(s) and titles once in a while to check for impersonations or piracy.⁴⁸³

202. A second mechanism of harm is that AI companies such as Meta provide LLMs that can create substitute but inferior works that will drive out authors’ IP from the market and diminish authors’ income. This creates what economists refer to as a “market for lemons.”⁴⁸⁴ Here, a lemon refers to a low-quality or inferior product. The problem of the market for lemons is widely taught in economics courses. As Charles Holt and Roger Sherman note, “[t]his lemons outcome is discussed in some introductory and most intermediate microeconomics courses, and it is examined in some detail in a range of applied courses: industrial organization, regulation, antitrust, managerial economics, law and economics, game theory and experimental economics.”⁴⁸⁵

203. The market-for-lemons problem is due to an asymmetry of information between suppliers of goods, who know the quality of their goods, and buyers of those goods, who do not know the quality of the

⁴⁷⁹ Cimerman, Agne and Danil Myakin. “AI-Generated Books Pose New Challenges for Authors.” *Metaverse Post* (Aug. 10, 2023). <<https://mpost.io/ai-generated-books-pose-new-challenges-for-authors>> (accessed Oct. 30, 2024).

⁴⁸⁰ *Id.*

⁴⁸¹ *Id.*

⁴⁸² Friedman, Jane. “IMHO: What Remedies Do Authors Have When Fraudulent Work Appears on Amazon?” *The Hot Sheet* (Aug. 16, 2023). <<https://hotsheetpub.com/2023/08/imho-what-remedies-do-authors-have-when-fraudulent-work-appears-on-amazon>> (accessed Oct. 30, 2024).

⁴⁸³ *Id.*

⁴⁸⁴ Akerlof, George A. “The Market for ‘Lemons’: Quality Uncertainty and Market Mechanism.” *Quarterly Journal of Economics* 84.3 (1970): 488-500.

⁴⁸⁵ Holt, Charles A. and Roger Sherman. “A Market for Lemons.” *Journal of Economic Perspectives* 13.1 (1999): 205-214 at 205.

goods. As Holt and Sherman explain: “If product quality cannot be observed by buyers prior to purchase, then sellers will be tempted to skimp on it. Buyers then become reluctant to pay high prices as they learn to expect low-quality products—or ‘lemons.’”⁴⁸⁶ Holt and Sherman point out that “the pressure of competition may cause quality to deteriorate to such low levels that the market may fail to exist.”⁴⁸⁷ The lemons problem arises when AI generated books are introduced in the marketplace because consumers may not know whether these books were generated by AI or a human author before purchasing and reading the book, and there may not be sufficient information available that accurately identifies AI generated works. Given the plethora of AI books imitating or summarizing specific human-published works, consumers may also mistakenly purchase AI books when they intended to purchase the author’s original work. It may be difficult to obtain third-party evaluations of these books because of the high volume of AI generated works and the greater speed and lower cost of production in comparison with authored books. Because it is difficult for buyers of books to observe the quality of the work in advance of purchase and consumption, it may be difficult for buyers to determine whether or not an AI LLM generated book is a sham or is inferior to a book written by a human author.

204. The Authors Guild sent an open letter signed by over 15,000 authors to Sam Altman, CEO, OpenAI; Sundar Pichai, CEO, Alphabet; Mark Zuckerberg, CEO, Meta; Emad Mostaque, CEO, Stability AI; Arvind Krishna, CEO, IBM; and Satya Nadella, CEO, Microsoft.⁴⁸⁸ The letter states that the usage of copyrighted works to train AI systems will generate content that will damage the writing profession: “As a result of embedding our writings in your systems, generative AI threatens to damage our profession by flooding the market with mediocre, machine-written books, stories, and journalism based on our work.”⁴⁸⁹

205. The Science Fiction and Fantasy Writers Association (“SFWA”) issued a letter regarding the difficulties with publishing stories in online and paper magazines due to AI LLMs generating substitute works. According to the SFWA, “[o]ver the last year, these venues, particularly the ones that pay higher rates for stories, have been inundated with AI-written stories. The editors uniformly report that these submissions are poorly conceived and written, far from being publishable, but the sheer volume materially interferes with the running of these magazines. Once submission systems are flooded with such content, it takes longer to read and reject a submission than it took someone to

⁴⁸⁶ *Id.*

⁴⁸⁷ *Id.*

⁴⁸⁸ “Open Letter to Generative AI Leaders.” *The Authors Guild*. <<https://actionnetwork.org/petitions/authors-guild-open-letter-to-generative-ai-leaders>> (accessed Nov. 4, 2024).

⁴⁸⁹ *Id.*

have an AI produce it in the first place. Every submitted work must be opened and considered to verify that the writers for whom the system was originally designed are not missed or forgotten.”⁴⁹⁰

206. The European Writers’ Council observes that “[f]or months, the global self-publishing provider Amazon has been flooded with bogus books by fake authors whose text and visual content have been mashed together by (re-)generative text and image output software. AI bots from click farms ‘read’ these nonsense works and pushed them into the bestseller lists. This led to a rapid decline in revenue for human authors by shared-revenue models, such as Kindle KDP (A pot of revenue divided by pages read and number of authors, similar to Spotify). At peak times, 80 out of 100 Kindle KDP bestsellers are AI editions.”⁴⁹¹ According to the European Writers’ Council, “[o]nly in September 2023 the retailer giant added a new section to its content guidelines for KDP focused on AI, which since then includes definitions of ‘AI-generated’, to label AI output.”⁴⁹²

207. In January 2024, however, journalist Sascha Brodsky reported that the AI-generated books trend was “[g]etting [w]orse[.]”⁴⁹³ According to Brodsky, “[t]he influx of AI-generated books on Amazon’s marketplace is creating major obstacles for human authors, experts say. It is also making it harder to distinguish between authentic authors and pseudonyms created by AI.”⁴⁹⁴ Brodsky further quotes an affected author as saying: “‘Whoever is doing this is obviously preying on writers who trust my name and think I have actually written these books. I have not. Most likely they have been generated by AI,’ author Jane Friedman said in a blog post after finding fake books supposedly written by her on Goodreads. ‘When I complained about this on Twitter/X, an author responded that she had to report 29 illegitimate books in just the last week alone. 29!’”⁴⁹⁵ Plagiarism Today states that Amazon’s control policies are failing: “Reports of AI books flooding the store continue, with one

⁴⁹⁰ Strauss, Victoria. “Artificial Intelligence and Copyright: SFWAs Comments to the US Copyright Office.” *Writer Beware* (Nov. 10, 2023). <<https://writerbeware.blog/2023/11/10/artificial-intelligence-and-copyright-sfwas-comments-to-the-us-copyright-office>> (accessed Nov. 4, 2024).

⁴⁹¹ “Analysis: The Success of Generative AI in the Book Sector Is Based on Theft.” *European Writers’ Council* (Sept. 28, 2023). <<https://europeanwriterscouncil.eu/gai-is-based-on-theft>> (accessed Oct. 1, 2024) (emphasis omitted). See also Creamer, Ella. “Self-Publishers Must Declare if Content Sold on Amazon’s Site Is AI-Generated.” *The Guardian* (Sept. 11, 2023). <<https://www.theguardian.com/books/2023/sep/11/self-publishers-must-declare-if-content-sold-on-amazons-site-is-ai-generated>> (accessed Nov. 4, 2024).

⁴⁹² “Analysis: The Success of Generative AI in the Book Sector Is Based on Theft.” *European Writers’ Council* (Sept. 28, 2023). <<https://europeanwriterscouncil.eu/gai-is-based-on-theft>> (accessed Oct. 1, 2024) (emphasis omitted). See also Creamer, Ella. “Self-Publishers Must Declare if Content Sold on Amazon’s Site Is AI-Generated.” *The Guardian* (Sept. 11, 2023). <<https://www.theguardian.com/books/2023/sep/11/self-publishers-must-declare-if-content-sold-on-amazons-site-is-ai-generated>> (accessed Nov. 4, 2024).

⁴⁹³ Brodsky, Sascha. “The AI-Generated Books Trend Is Getting Worse.” *AI Business* (Jan. 24, 2024). <<https://aibusiness.com/responsible-ai/the-ai-generated-books-phenomenon-is-getting-worse>> (accessed Nov. 4, 2024).

⁴⁹⁴ *Id.*

⁴⁹⁵ *Id.*

reporter saying that Amazon promotes the books via Kindle ads. Perhaps most worrisome, many of these ads are for AI-generated children’s books.”⁴⁹⁶

208. Third, AI companies such as Meta provide LLMs that can create similar or sham works that will damage the reputation of authors and thus diminish the market for their work. According to a study in the *California Management Review*, “as AI becomes more prevalent, the number of cases where its application violates social norms and values rises.”⁴⁹⁷ The European Writers’ Council states that “AI harms human authors, their income and reputation through fake authors, fake books, fake readers – and identity theft[.]”⁴⁹⁸

209. Fourth, because AI companies such as Meta provide LLMs that can create substitute works that flood the market, authors would need to incur prohibitive marketing costs to sell their works. Joanna Penn points out that the speed and low cost of AI-generated content will compete with authors: “If content is produced by AI at a faster rate in every medium — and possibly with the addition of translated material — then the tsunami of content will soon bury us all.”⁴⁹⁹ Penn adds authors will be harmed further because they will not be able to bear the additional cost of marketing their work: “Authors and publishers already have to pay for Amazon Advertising to even be seen in the Amazon stores right now and this model will soon be broken as ads become too expensive for most.”⁵⁰⁰

210. These four major effects demonstrate that the usage of copyrighted works to train AI systems will diminish or eliminate the incentives for writers to produce creative works. The Authors Guild letter states: “In the past decade or so, authors have experienced a forty percent decline in income, and the current median income for full-time writers in 2022 was only \$20,000. The introduction of AI threatens to tip the scale to make it even more difficult, if not impossible, for writers—especially

⁴⁹⁶ Bailey, Jonathan. “Why Amazon Is Overrun with Plagiarism and AI Garbage.” *Plagiarism Today* (Apr. 17, 2024). <<https://www.plagiarismtoday.com/2024/04/17/why-amazon-is-overrun-with-plagiarism-and-ai-garbage>> (accessed Nov. 4, 2024).

⁴⁹⁷ Holweg, Matthias, et al. “The Reputational Risks of AI.” *California Management Review* (Jan. 24, 2022). <<https://cmr.berkeley.edu/2022/01/the-reputational-risks-of-ai>> (accessed Nov. 4, 2024).

⁴⁹⁸ “Analysis: The Success of Generative AI in the Book Sector Is Based on Theft.” *European Writers’ Council* (Sept. 28, 2023). <<https://europeanwriterscouncil.eu/gai-is-based-on-theft>> (accessed Oct. 1, 2024) (emphasis omitted).

⁴⁹⁹ Penn, Joanna. “9 Ways that Artificial Intelligence (AI) Will Disrupt Authors and the Publishing Industry.” *The Creative Penn* (July 1, 2019). <<https://www.thecreativepenn.com/2019/07/01/ai-disruption-publishing-authors>> (accessed Nov. 4, 2024).

⁵⁰⁰ *Id.*

young writers and voices from under-represented communities—to earn a living from their profession.”⁵⁰¹

211. The European Guild for Artificial Intelligence Regulation, in a manifesto signed by over 8,000 people, raises similar concerns about data used to train AI LLMs: “This exploitation of our work and data not only does not respect the rights that regulate our society, presenting a huge security risk: it is also severely damaging the art market, potentially scarring it forever [sic]. We see this as only the beginning of a crisis that will afflict all sorts of jobs and occupation, whether they are creative jobs or not. The art market is the first one to be affected only because of its structural vulnerabilities, which make it an easy prey.”⁵⁰²
212. If fiction and nonfiction books can be used as training data for LLMs, resulting in works that compete with their creative works, there will be a significant impact on the production of creative works. AI construction of fiction or nonfiction works that diminishes demand for related creative work by authors will discourage authors from producing those works. Not only will this harm authors, but it will also harm their readers. Without the production of creative works by authors, the data-driven AI production will itself suffer due to a lack of new human-generated training data.
213. AI LLM generated substitute works have raised concerns that AI could completely displace creators of copyrighted content including writers and artists.⁵⁰³ Former OpenAI employee Suchir Balaji left the company over concerns that its LLMs’ use of copyrighted training data “violated the law and that

⁵⁰¹ “Open Letter to Generative AI Leaders.” *The Authors Guild*. <<https://actionnetwork.org/petitions/authors-guild-open-letter-to-generative-ai-leaders>> (accessed Nov. 4, 2024).

⁵⁰² “Our Manifesto for AI Companies Regulation in Europe.” *European Guild for Artificial Intelligence Regulation* (Nov. 4, 2023). <<https://www.egair.eu/#manifesto>> (accessed Nov. 4, 2024).

⁵⁰³ See, e.g., Bilton, Nick. “The New Generation of A.I. Apps Could Make Writers and Artists Obsolete.” *Vanity Fair* (June 2, 2022). <<https://www.vanityfair.com/news/2022/06/the-new-generation-of-ai-apps-could-make-writers-and-artists-obsolete>> (accessed Oct. 28, 2024) (“Over the past few months, new advancements in A.I. have made it clear that writers, illustrators, photographers, journalists, and novelists could soon be driven from the workforce and replaced by high-tech player pianos.”); Soto, Christopher. “Will AI Replace Writers – and the Rest of Us? An L.A. Scholar-Poet Games Out the Future.” *Los Angeles Times* (Aug. 23, 2023). <<https://www.latimes.com/entertainment-arts/books/story/2023-08-23/will-ai-replace-writers-christopher-soto-jackie-wang-game-out-the-future>> (accessed Oct. 28, 2024) (“We may soon reach a point where certain types of writing — screenwriting, journalism, web content — and certain para-literary activities — editing, proofreading, researching — could be fully or partially automated. ... Should AI be allowed to imitate living writers and artists, and will the imitations be commercialized at the expense of living creators?”); Martin, Tiffany Yates. “Will AI Replace Writers? It Already Is.” *FoxPrint Editorial* (May 11, 2023). <<https://foxprinteditorial.com/2023/05/11/will-ai-replace-writers-it-already-is/>> (accessed Oct. 28, 2024); Clark, Elijah. “The End of Originality: Is AI Replacing Real Artists?” *Forbes* (Dec. 23, 2023). <<https://www.forbes.com/sites/elijahclark/2023/12/23/the-end-of-originality-is-ai-replacing-real-artists/>> (accessed Oct. 28, 2024) (“AI is not just showcasing its capabilities; it’s actively participating in the creative industry. It’s creating beautiful artwork, designing innovative products, composing music and even writing award-winning novels.”).

technologies like ChatGPT were damaging the internet.”⁵⁰⁴ He observed that LLMs like ChatGPT (and Llama) were “designed to imitate online data ... so they can substitute for ‘basically anything’ on the internet, from news stories to online forums.”⁵⁰⁵ This all has the effect that, in his opinion, “ChatGPT and other chatbots ... are destroying the commercial viability of the individuals, business and internet services that created the digital data used to train these A.I. systems.”⁵⁰⁶ An article in *Vanity Fair* suggests that “writers, illustrators, photographers, journalists, and novelists could soon be driven from the workforce and replaced by high-tech player pianos” in the form of AI and LLMs.⁵⁰⁷

214. These observations imply that Llama-generated substitute works can be expected to diminish the market for authors’ works. To evaluate the potential economic harm from AI LLMs displacing authors’ works, it is useful to review the U.S. and global markets for fiction and nonfiction books. Total revenues for the U.S. book publishing industry were between \$25 billion and \$26 billion from 2016 to 2020,⁵⁰⁸ increasing to \$29.33 billion in 2021 and \$28.10 billion in 2022.⁵⁰⁹ For 2023, the Association of American Publishers estimated \$29.9 billion in total publishing revenue, of which \$18.7 billion were sales of trade books.⁵¹⁰

⁵⁰⁴ Metz, Cade. “Former OpenAI Researcher Says the Company Broke Copyright Law.” *The New York Times* (Oct. 23, 2024). <<https://www.nytimes.com/2024/10/23/technology/openai-copyright-law.html>> (accessed Oct. 25, 2024).

⁵⁰⁵ *Id.*

⁵⁰⁶ *Id.*

⁵⁰⁷ Bilton, Nick. “The New Generation of A.I. Apps Could Make Writers and Artists Obsolete.” *Vanity Fair* (June 2, 2022). <<https://www.vanityfair.com/news/2022/06/the-new-generation-of-ai-apps-could-make-writers-and-artists-obsolete>> (accessed Oct. 28, 2024).

⁵⁰⁸ Rowe, Adam. “5 Facts About the \$25.7 Billion That U.S. Publishers Earned in 2020.” *Forbes* (Dec. 29, 2021). <<https://www.forbes.com/sites/adamrowe/2021/12/29/5-facts-about-the-257-billion-that-us-publishers-earned-in-2020/?sh=3b832ca9179d>> (accessed Oct. 29, 2024).

⁵⁰⁹ “AAP StatShot Annual Report For 2021: Book Publishing Revenues Up 12.3% For the Year, Reaching All-Time High of \$29.33 Billion.” *Association of American Publishers* (Sept. 16, 2022). <<https://publishers.org/news/aap-statshot-annual-report-for-2021-book-publishing-revenues-up-12-3-for-the-year-reaching-all-time-high-of-29-33-billion>> (accessed Oct. 3, 2024); “AAP StatShot Annual Report: Publishing Revenues Totaled \$28.10 Billion for 2022.” *Association of American Publishers* (May 31, 2023). <<https://publishers.org/news/aap-statshot-annual-report-publishing-revenues-totaled-28-10-billion-for-2022>> (accessed Oct. 3, 2024).

⁵¹⁰ Trade books are any books published and sold for consumers or general readership. “Trade Book Definition.” *Britannica*. <<https://www.britannica.com/dictionary/trade-book>> (accessed Oct. 3, 2024); “AAP StatShot Annual Report: Publishing Revenues Totaled \$29.9 Billion for 2023.” *American Association of Publishers* (Aug. 22, 2024). <<https://publishers.org/news/aap-statshot-annual-report-publishing-revenues-totaled-29-9-billion-for-2023>> (accessed Oct. 3, 2024).

215. At the global level, WIPO has estimated a total global publishing revenue of \$76.3 billion in 2022.⁵¹¹

Table 4: *Total Publishing Industry Revenue by Country (USD million), 2022* shows 2022 publishing revenues for select countries by publishing sector.

Table 4: Total Publishing Industry Revenue by Country (USD million), 2022⁵¹²

Country	Total	Trade	Educational	Share of Total (%)	
				Trade	Educational
Austria	369.1	314.2	54.9	85.1	14.9
Azerbaijan (a)	17.8	17.8	-	-	-
Belgium (b)	318.1	135.9	182.2	42.7	57.3
Brazil	1,113.4	458.0	655.4	41.1	58.9
Colombia	211.5	105.2	106.3	49.8	50.2
Czech Republic	153.0	119.0	34.1	77.7	22.3
Denmark	265.7	110.4	155.3	41.5	58.5
Finland	290.3	133.9	156.3	46.1	53.9
France	2,909.1	2,156.4	752.7	74.1	25.9
Germany	9,945.0	7,995.8	1,949.2	80.4	19.6
Hungary (a)	125.7	125.7	-	-	-
Iceland	33.5	-	-	-	-
India (c, d)	9,131.5	342.2	8,517.3	3.7	93.3
Ireland (c)	30.7	20.1	10.6	65.6	34.4
Italy	3,567.8	-	-	-	-
Japan	9,315.7	8,686.1	629.7	93.2	6.8
Malta	4.0	2.6	1.4	65.0	35.0
Mexico *	467.0	124.5	342.4	26.7	73.3
Netherlands (Kingdom of the)	1,006.0	341.9	664.1	34.0	66.0
New Zealand	111.7	89.0	22.7	79.7	20.3
Norway	337.1	163.8	173.2	48.6	51.4
Philippines (a, c)	380.6	380.6	-	-	-
Portugal	300.1	184.3	115.8	61.4	38.6
Republic of Korea *	6,654.3	3,390.8	3,263.5	51.0	49.0
Russian Federation	1,486.8	1,251.7	235.1	84.2	15.8
Spain	2,862.7	1,619.8	1,242.9	56.6	43.4
Sweden (a)	199.9	199.9	-	-	-

⁵¹¹ WIPO collects publishing industry data provided by 35 countries through national publishers' associations and copyright authorities. The data is limited to trade and educational sectors with issued book identifiers, such as ISBN or DOI. "The Global Publishing Industry in 2022." *World Intellectual Property Organization* (2023): 1-33 at 6-8; "World Intellectual Property Organization." *United States Patent and Trademark Office*. <<https://www.uspto.gov/ip-policy/patent-policy/world-intellectual-property-organization>> (accessed Oct. 3, 2024).

⁵¹² WIPO indicated certain data characteristics, represented by the following markings: (*) 2021 data, (a) trade sector only, (b) French-speaking region, (c) print format only, (d) USD 272 million (3% of total) are unclassified. "The Global Publishing Industry in 2022." *World Intellectual Property Organization* (2023): 1-33 at 8.

Türkiye	638.3	169.1	469.2	26.5	73.5
UK	5,024.3	2,786.1	2,238.1	55.5	44.5
US	26,150.9	17,360.8	8,790.1	66.4	33.6

216. WIPO also reported revenues specifically for fiction and non-fiction books for ten countries with large book markets. In 2022, total non-fiction revenue across these ten countries was approximately \$3.3 billion, and total fiction revenue was \$2.1 billion.⁵¹³ See **Table 5: Total Copies Sold 2019-2022** and **Table 6: Sales Revenue, 2019-2022**.

Table 5: Total Copies Sold, 2019-2022⁵¹⁴

Country	2019	2020	2021	2022	2022 Distribution (%)		
					Fiction	Children's	Non-fiction
Australia	60.6	67.2	65.5	70.9	25.9	45.8	28.1
Brazil	41.5	43.4	56.9	58.6	29.7	24.0	45.9
India	30.9	29.7	38.2	35.9	15.6	28.8	53.5
Ireland	12.2	13.1	13.3	13.4	28.9	36.1	34.5
Italy	93.1	90.8	109.3	107.0	34.6	22.8	42.6
Mexico	15.7	13.8	17.6	20.3	20.0	30.1	42.4
New Zealand	6.4	6.4	6.8	6.6	22.5	46.4	30.9
South Africa	8.6	7.0	8.0	8.2	22.6	35.3	40.8
Spain	62.6	55.4	65.1	66.5	33.2	31.8	24.0
United Kingdom (a)	191.4	202.0	212.6	208.8	29.8	35.3	34.7

⁵¹³ For each country, I calculated the revenue for fiction (non-fiction) books as the product of revenue from books sold in 2022 and the fraction of revenue attributed to fiction (non-fiction) books, as reported in Table 5. See *id.* at 20.

⁵¹⁴ *Id.*

Table 6: Sales Revenue, 2019-2022⁵¹⁵

Country	2019	2020	2021	2022	2022 Distribution (%)		
					Fiction	Children's	Non-fiction
Australia	793	862	950	940	26.8	29.4	43.5
Brazil	443	349	435	493	27.1	21.0	51.5
India	138	138	174	158	14.3	24.9	58.9
Ireland	165	184	196	179	27.4	28.3	43.7
Italy	1,619	1,634	1,970	1,717	34.5	18.1	47.4
Mexico	174	145	209	256	21.4	24.6	46.6
New Zealand	93	91	108	94	25.5	29.4	44.9
South Africa	106	81	109	108	23.9	25.1	49.9
Spain	1,095	997	1,200	1,115	32.4	24.8	27.3
United Kingdom (a)	2,127	2,256	2,506	2,229	26.6	24.5	48.7

217. The five largest trade book publishers in the U.S., known as the “Big Five,” control 90 percent of the market.⁵¹⁶ These five publishers are Penguin Random House, Simon & Schuster, Hachette Book Group, HarperCollins Publishers, and Macmillan Publishing Group, LLC.⁵¹⁷ The Big Five have “grown concertedly through acquisition,” led by Penguin Random House with six acquisitions.⁵¹⁸ Combined, the publicly reported global revenues for the Big Five⁵¹⁹ was at least \$10.5 billion in 2022. See **Figure 10: Book Publishing Global Revenue by Publisher.**

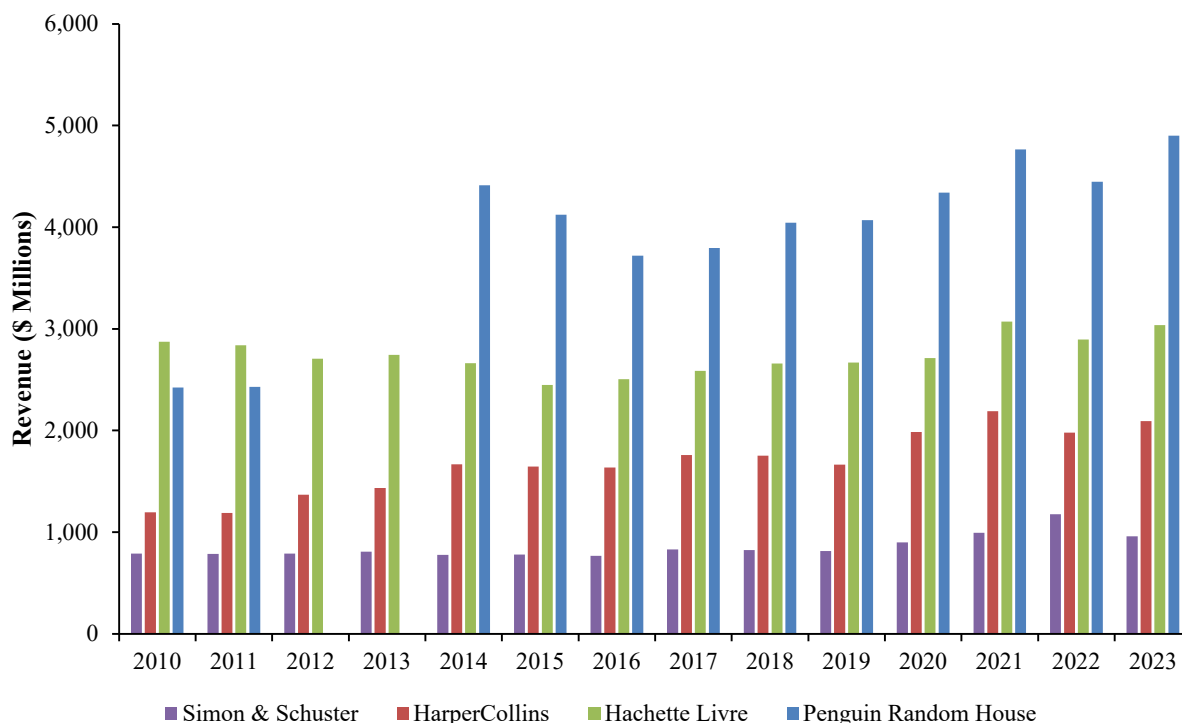
⁵¹⁵ *Id.*

⁵¹⁶ Shepardson, David. “U.S. Asks Judge to Block Book Publisher Merger as Trial Closes.” *Reuters* (Aug. 19, 2022). <<https://www.reuters.com/markets/deals/us-asks-judge-block-book-publisher-merger-trial-closes-2022-08-19>> (accessed Oct. 3, 2024); See Lamb, David. “Trends In U.S. Trade Book Publisher Mergers and Acquisitions.” *Publishing Research Quarterly* 36 (2020): 323-334 at 328.

⁵¹⁷ Milliot, Jim. “Over the Past 25 Years, the Big Publishers Got Bigger – and Fewer.” *Publishers Weekly* (Apr. 19, 2022). <<https://www.publishersweekly.com/pw/by-topic/industry-news/publisher-news/article/89038-over-the-past-25-years-the-big-publishers-got-bigger-and-fewer.html>> (accessed Oct. 3, 2024).

⁵¹⁸ Lamb, David. “Trends In U.S. Trade Book Publisher Mergers and Acquisitions.” *Publishing Research Quarterly* 36 (2020): 323-334 at 328.

⁵¹⁹ This includes Penguin Random House, Hachette Livre, HarperCollins, and Simon & Schuster, and excludes Macmillan Publishing Group, LLC as it is a private company and does not disclose financial data.

Figure 10: Book Publishing Global Revenue by Publisher^{520, 521}

218. Penguin Random House,⁵²² the largest publishing house by revenue, publishes over 70,000 digital and 15,000 print titles across their three hundred imprints annually.⁵²³ In 2023, Penguin Random

⁵²⁰ Bloomberg Intelligence reports annual revenues for largest public publishing companies in North America on their Publishing and Broadcasting dashboard (BI LOCLN). Bloomberg, L.P. (accessed Jan. 9, 2025); Simon & Schuster revenue from 2019-2023 was extracted from Paramount Global's (formerly ViacomCBS, Inc.) Form 10-K, where it was presented as a discontinued operation due to intention to sell the subsidiary. ViacomCBS, Inc. *Form 10-K* (Dec. 31, 2019) at II-26; ViacomCBS, Inc. *Form 10-K* (Dec. 31, 2020) at II-17; ViacomCBS, Inc. *Form 10-K* (Dec. 31, 2021) at II-14; Paramount Global. *Form 10-K* (Dec. 31, 2022) at II-67; Paramount Global. *Form 10-K* (Dec. 31, 2023) at II-14.

⁵²¹ Original data labels included the parent company in addition to the subsidiary, or only the parent company. The labels have been changed to align with the subsidiary.

⁵²² Random House and Penguin merged in 2013, creating the largest publishing house in the world, accounting for the large revenue increase in 2014 shown in Figure 10. Bertelsmann, originally the parent company of Random House, held majority ownership of the merged publishing house until acquiring the remaining stake in 2019. The data in Figure 10 prior to 2013 regarding Penguin Random House contains only Random House figures reported by Bertelsmann. See Frieden, Terry. "U.S. OKs Penguin-Random House Merger." *CNN* (Feb. 14, 2013). <<https://www.cnn.com/2013/02/14/business/penguin-random-house-merger/index.html>> (accessed Oct. 3, 2024); "Bertelsmann Acquires Full Ownership of Penguin Random House." *Penguin Random House* (Dec. 18, 2019). <<https://global.penguinrandomhouse.com/announcements/bertelsmann-acquires-full-ownership-of-penguin-random-house>> (accessed Oct. 3, 2024).

⁵²³ "Imprints." *Penguin Random House*. <<https://www.penguinrandomhouse.com/imprints>> (accessed Oct. 3, 2024).

House's global revenue was \$4.9 billion, with 1.6 percent attributed to "rights and licensing."⁵²⁴ From 2017-2022, this rights and licensing accounted for between 1.3 and 1.7 percent of revenue.⁵²⁵

219. HarperCollins claims to be the second largest consumer book publisher in the world by revenue, owning 120 imprints that publish approximately 10,000 new titles per year of both fiction and nonfiction.⁵²⁶ In 2023, HarperCollins reported a global revenue of \$2.09 billion.⁵²⁷

220. Hachette Livre,⁵²⁸ parent of U.S. subsidiary Hachette Book Group, claims to be the world's third-largest trade publisher, producing 15,000 titles per year and a revenue of \$2.8 billion.⁵²⁹ Hachette Livre is comprised of two hundred global imprints. Hachette Book Group's 43 U.S. imprints publish 2,100 adult trade books and 500 children's and young adult trade books.⁵³⁰

221. Simon & Schuster is an independent company that publishes over 2,000 titles annually.⁵³¹ Previously, the company was a subsidiary of Paramount Global, formerly known as Viacom CBS, until sold to a private equity firm for \$1.62 billion in 2023.⁵³² In 2023, Paramount Global reported that Simon & Schuster's revenue was \$958 million.⁵³³

222. Macmillan is part of privately held German media company Holtzbrinck Publishing Group, with imprints publishing multiple genres, including non-fiction and fiction titles.⁵³⁴

⁵²⁴ Bloomberg, L.P. (accessed Nov. 7, 2024); "Annual Report 2023." *Bertelsmann* (2023) at 28.

⁵²⁵ "Annual Report 2022." *Bertelsmann* (2022) at 25; "Annual Report 2021." *Bertelsmann* (2021) at 21; "Annual Report 2020." *Bertelsmann* (2020) at 21; "Annual Report 2019." *Bertelsmann* (2019) at 20; "Annual Report 2018." *Bertelsmann* (2018) at 19; "Annual Report 2017." *Bertelsmann* (2017) at 19.

⁵²⁶ "Annual Report 2022." *NewsCorp* (2022) at 8; "About Us." *HarperCollins Publishers*.

<<https://www.harpercollins.com/pages/aboutus>> (accessed Oct. 3, 2024).

⁵²⁷ Bloomberg, L.P. (accessed Nov. 7, 2024).

⁵²⁸ Hachette Livre is the global publishing group under Lagardère, and Hachette Book Group is the U.S. subsidiary of Hachette Livre. Revenue data in Figure 10 is global book publishing revenue reported under Hachette Livre, including revenue from Hachette Book Group.

⁵²⁹ "The Hachette Livre Group: An Overview." *Hachette Livre*. <<https://www.hachette.com/en/hachette-livre-group-overview>> (accessed Oct. 3, 2024).

⁵³⁰ "Book Publishing." *Hachette Livre*. <<https://www.hachette.com/en/book-publishing>> (accessed Oct. 3, 2024); "The Hachette Livre Group: An Overview." *Hachette Livre*. <<https://www.hachette.com/en/hachette-livre-group-overview>> (accessed Oct. 3, 2024).

⁵³¹ "Corporate Overview." *Simon & Schuster*. <<https://about.simonandschuster.biz/corporate-overview>> (accessed Oct. 3, 2024).

⁵³² Blair, Elizabeth. "Paramount Sells Simon & Schuster to Private Investment Firm." *NPR* (Aug. 7, 2023). <<https://www.npr.org/2023/08/07/1192539219/simon-schuster-sold>> (accessed Oct. 3, 2024).

⁵³³ Paramount Global. *Form 10-K* (Dec. 31, 2023) at II-14.

⁵³⁴ "Macmillan Publishers is a Leading Publishing Company and Home to Some of the World's Most Cherished Authors and Creators." *Macmillan Publishers*. <<https://us.macmillan.com/about>> (accessed Oct. 3, 2024).

223. This review of the size and significance of the U.S. and global markets for fiction and nonfiction books shows the devastating potential negative economic impact of AI LLM generated substitute works on the market for authors' works. This economic analysis of the publishing industry provides support for the observation that unauthorized and uncompensated usage of copyrighted works by Meta's AI LLM program causes harm to authors.
224. Analyzing AI's real-world impact on usage rates and revenue for other companies and industries can help provide a quantitative estimate of the economic impact of AI LLM generated substitute works on the publishing industry. For example, online education company Chegg stated that its business was suffering as students were increasingly using the LLM ChatGPT as a study aid instead of paying for Chegg's products and tools,⁵³⁵ with Chegg CEO Dan Rosensweig saying, "since March we saw a significant spike in student interest in ChatGPT. We now believe it's having an impact on our new customer growth rate."⁵³⁶ Shares of Chegg plummeted roughly 48 percent on the announcement,⁵³⁷ and as of April 2024 the company's shares had tumbled 95 percent from their peak in 2021, coincident with year-over-year subscriber declines since ChatGPT's release.⁵³⁸ According to a report, "Chegg's 2024 first quarter revenue of \$174.4 million was down 7% compared to the same period a year prior and down nearly 14% from the first quarter of 2022."⁵³⁹ Chegg's year-over-year subscriber decline is shown in **Figure 11: Chegg Subscriber Growth, Q3 2021-Q3 2024**.

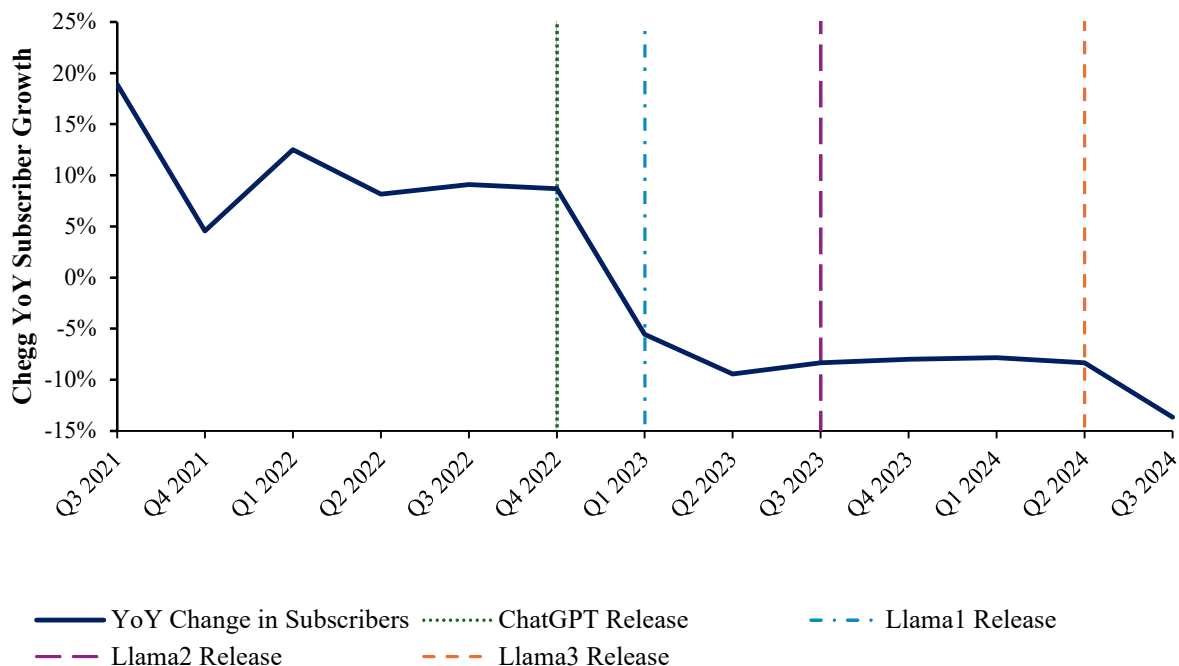
⁵³⁵ Min, Sarah. "Chegg Shares Drop More than 40% After Company Says ChatGPT Is Killing Its Business." *CNBC* (May 2, 2023). <<https://www.cnbc.com/2023/05/02/chegg-drops-more-than-40percent-after-saying-chatgpt-is-killing-its-business.html>> (accessed Oct. 28, 2024).

⁵³⁶ *Id.*

⁵³⁷ *Id.*

⁵³⁸ Schafer, Josh. "Chegg Stock Crashes as Free AI Tools Send Online Education Company 'Spiraling.'" *Yahoo! Finance* (Apr. 30, 2024). <<https://finance.yahoo.com/news/chegg-stock-crashes-as-free-ai-tools-send-online-education-company-spiraling-135522894.html>> (accessed Oct. 28, 2024).

⁵³⁹ *Id.*

Figure 11: Chegg Subscriber Growth, Q3 2021-Q3 2024⁵⁴⁰

225. The chief revenue officer of Textbroker, a content platform providing “credible, human-written copy on a huge range of topics,” remarked to the Washington Post that “People are trying automated content right now, and so that has slowed down our growth[.]”⁵⁴¹

226. Stack Overflow, a question-and-answer site focused on coding and programming, provides another example. Stack Overflow is designed for developers to seek help, share knowledge, and collaborate on technical problems. The platform is structured around user-generated content where questions are posed, and community members provide solutions, often with a voting system that highlights the

⁵⁴⁰ Chegg, Inc. *Form 8-K* (Oct. 26, 2020) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (Feb. 8, 2021) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (May 3, 2021) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (Aug. 9, 2021) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (Nov. 1, 2021) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (Feb. 7, 2022) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (May 2, 2022) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (Aug. 4, 2022) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (Nov. 1, 2022) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (Feb. 6, 2023) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (May 1, 2023) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (Aug. 7, 2023) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (Oct. 30, 2023) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (Feb. 5, 2024) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (Apr. 29, 2024) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (Aug. 5, 2024) Exhibit 99.1 at 1; Chegg, Inc. *Form 8-K* (Nov. 12, 2024) Exhibit 99.1 at 1.

⁵⁴¹ Oremus, Will. “He Wrote a Book on a Rare Subject. Then a ChatGPT Replica Appeared on Amazon.” *The Washington Post* (May 5, 2023). <<https://www.washingtonpost.com/technology/2023/05/05/ai-spam-websites-books-chatgpt/>> (accessed Oct. 30, 2024).

most useful answers. Stack Overflow fosters a community where expertise is shared, and users can build reputations based on the quality of their contributions.⁵⁴²

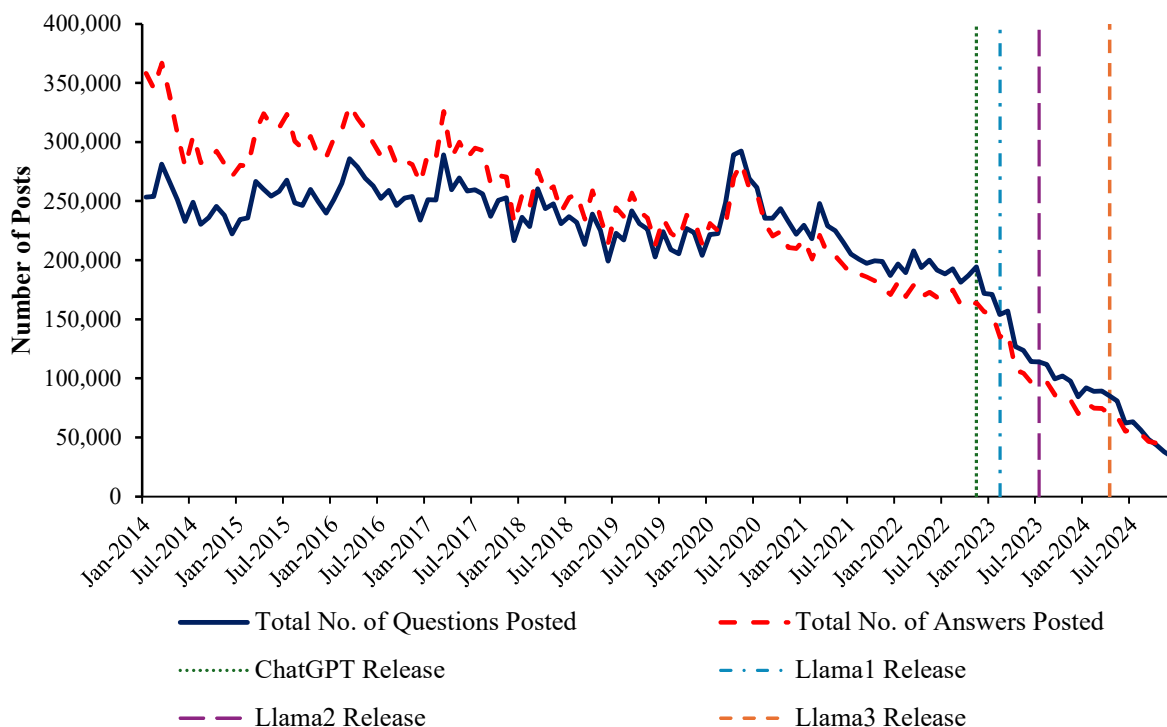
227. LLMs like Llama had a significant substitution effect on online Q&A communities.⁵⁴³ In the realm of online Q&A communities, platforms like Stack Overflow traditionally thrived on the collaborative exchange of knowledge, where users contribute questions and answers to build a repository of information. However, with LLMs' ability to provide immediate, AI-generated responses, the dynamics of these communities changed. Users seeking quick answers to technical or general questions could bypass these platforms entirely, opting instead for the convenience of AI. I used daily data on total questions and answers posted on Stack Overflow to analyze the impact of ChatGPT on user activity.⁵⁴⁴ As shown below in **Figure 12: Questions and Answers Posted on Stack Overflow, 2014-2024** the decline in the year-over-year questions and answers posted on Stack Overflow is significantly steeper after the release of ChatGPT in November 2022⁵⁴⁵ and the Llama models throughout 2023 when compared to the prior activity.

⁵⁴² "About Stack Overflow." *Stack Overflow*. <<https://stackoverflow.co/company/press>> (accessed Nov. 1, 2024).

⁵⁴³ Burch, Gordon, Dokyun Lee, and Zhichen Chen. "The Consequences of Generative AI for Online Knowledge Communities." *Scientific Reports* 14.1 (2024): 1-8.

⁵⁴⁴ Data on Stack Overflow questions and answers was obtained via the Stack Exchange Data Explorer for the period January 1, 2014, to December 31, 2024. "Stack Exchange Data Explorer." *Stack Exchange*. <<https://data.stackexchange.com/stackoverflow/query/new>> (accessed Jan. 9, 2025).

⁵⁴⁵ "Introducing ChatGPT." *OpenAI*. <<https://openai.com/index/chatgpt>> (accessed Dec. 4, 2024).

Figure 12: Questions and Answers Posted on Stack Overflow, 2014-2024

228. Although the number of questions and answers posted on Stack Overflow had been declining prior to the release of ChatGPT, this decline accelerated significantly after the release of ChatGPT in November 2022. The average year-over-year change in number of questions posted during the twenty-five months prior to the release of ChatGPT was -10.48 percent, compared to an average -40.77 percent decline for the same duration after the release.⁵⁴⁶

229. A study by Hui et al. found that “after ChatGPT’s release, the number of monthly jobs for writing-related freelancers on Upwork declined by 2%, while monthly earnings declined by 5.2%.”⁵⁴⁷ The study by Hui et al. also found that “[a] similar impact hit image-related workers (designers, image editors and artists) after the releases of two image-based AI tools, DALL-E, in April 2022, and Midjourney, in July 2022. Workers in those fields saw even steeper declines—a 3.7% drop in

⁵⁴⁶ November 30 was excluded from this analysis since it is the date of ChatGPT’s release.

⁵⁴⁷ Koziatsek, Suzanne. “Study: AI Tools Cause a Decline in Freelance Work and Income—at Least in the Short Run.” *Washington University Olin Business School* (Aug. 24, 2023). <<https://olin.wustl.edu/about/news-and-media/news/2023/08/study-ai-tools-cause-a-decline-in-freelance-work-and-income-at-least-in-the-short-run.php>> (accessed Jan. 5, 2025); Hui, Xiang, et al. “The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market.” *Washington University in St. Louis and New York University* (2023): 1-30.

monthly jobs and a 9.4% loss of income.”⁵⁴⁸ For 2023, the Association of American Publishers estimated \$29.9 billion in total publishing revenue.⁵⁴⁹ Publishing revenue fell by about 0.8 percent in 2023 with a decline in unit sales of 5.7 percent.⁵⁵⁰ An Authors Guild survey of over 5,000 authors reported that the median income for full-time authors from their books in 2022 was \$10,000.⁵⁵¹

230. As noted above and shown in Figures 11 and 12, Chegg’s usage declined by about 14 percent and Stack Overflow experienced an additional decline of over 30 percent. The decline in income for freelancers on Upwork was over 5 percent and for image-related workers was over 9 percent. If the decline in revenue due to AI equals the low end of these figures, this represents a potential loss of revenue for publishers of about \$1.5 billion per year and for authors of about \$500 per year per author on average.⁵⁵² If the decline in revenue due to AI equals the high end of these figures at 30 percent, publishers could lose roughly \$9 billion in revenue per year and individual authors could lose on average \$3,000 per year.⁵⁵³

IX. DETERMINING LICENSING ROYALTIES FOR AI LLM TRAINING DATA

231. In this section, I provide an economic analysis that explains why comparable licenses for AI LLM training data give an estimate of the market value of AI LLM training data. The market value of comparable licenses for AI LLM training data provides an indication of the benefits to buyers and the costs to sellers of AI LLM training data. Comparable licenses for AI LLM training data are useful as indicators of market value because they provide an indication of an agreement between a willing buyer and a willing seller. Comparable licenses for AI LLM training data reflect the necessity of data in training AI LLMs.

⁵⁴⁸ Koziatsek, Suzanne. “Study: AI Tools Cause a Decline in Freelance Work and Income—at Least in the Short Run.” *Washington University Olin Business School* (Aug. 24, 2023). <<https://olin.wustl.edu/about/news-and-media/news/2023/08/study-ai-tools-cause-a-decline-in-freelance-work-and-income-at-least-in-the-short-run.php>> (accessed Jan. 5, 2025); Hui, Xiang, et al. “The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market.” *Washington University in St. Louis and New York University* (2023): 1-30.

⁵⁴⁹ Milliot, Jim. “Publishing Revenue Fell Slightly in 2023, but Unit Sales Dropped 5.7%.” *Publishers Weekly* (Aug 22, 2024). <<https://www.publishersweekly.com/pw/by-topic/industry-news/financial-reporting/article/95788-publishing-industry-sales-fell-slightly-in-2023.html>> (accessed Jan. 5, 2025) (“The final publishing industry sales estimates from the Association of American Publishers for 2023 are in, and indicate that sales in the year fell 0.8% from 2022, to \$29.9 billion.”).

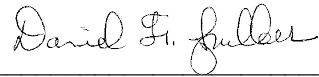
⁵⁵⁰ *Id.* (“While dollar sales fell by only 0.8%, unit sales were down 5.7%”).

⁵⁵¹ “Key Takeaways from the Authors Guild’s 2023 Author Income Survey.” *The Authors Guild* (Sept. 27, 2023). <<https://authorsguild.org/news/key-takeaways-from-2023-author-income-survey>> (accessed Jan. 8, 2025).

⁵⁵² That is, \$29.9 billion * 0.05 = \$1.495 billion, and \$10,000 * 0.05 = \$500.

⁵⁵³ \$29.9 billion * 0.3 = \$8.97 billion, and \$10,000 * 0.3 = \$3,000.

Dated: January 10, 2025

A handwritten signature in cursive script, reading "Daniel F. Spulber". The signature is written in black ink and is positioned above a horizontal line.

Daniel F. Spulber